

03/05/99



13-408 U.S. PTO

# UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.

862.2720

First Named Inventor or Application Identifier

MASAYUKI YAMADA

Express Mail Label No.

## APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

## ADDRESS TO:

Assistant Commissioner for Utility Patents  
Box Patent Application  
Washington, DC 202311. ☒ Fee Transmittal Form  
(Submit an original, and a duplicate for fee processing)2. ☒ Specification Total Pages **29**3. ☒ Drawing(s) (35 USC 113) Total Sheets **9**4. ☒ Oath or Declaration Total Pages **2**a. ☒ Newly executed (original or copy)b. ☐ Unexecuted for information purposesc. ☐ Copy from a prior application (37 CFR 1.63(d))  
(for continuation/divisional with Box 17 completed)  
(Note Box 5 below)i. ☐ DELETION OF INVENTOR(S)Signed Statement attached deleting inventor(s) named in  
the prior application, see 37 CFR 1.63(h)(2) and 1.39(b).5. ☐ Incorporation By Reference (see Box 4c is checked)  
The entire disclosure of the prior application, from which a copy of the oath or  
declaration is supplied under Box 4c, is considered as being part of the disclosure of  
the accompanying application and is hereby incorporated by reference therein.6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and/or Amino Acid Sequence Submission  
(if applicable, all necessary)a. ☐ Computer Readable Copyb. ☐ Paper Copy (identical to computer copy)c. ☐ Statement verifying identity of above copies

## ACCOMPANYING APPLICATION PARTS

8. ☒ Assignment Papers (cover sheet & documents)9. ☐ 37 CFR 3.73(b) Statement (when there is an assignee) ☐ Power of Attorney10. ☐ English Translation Document (if applicable)11. ☐ Information Disclosure Statement (IDS/PTO-1449) ☐ Copies of IDS Citations12. ☐ Preliminary Amendment13. ☒ Return Receipt Postcard (MPEP 503)  
(Should be specifically itemized)14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application  
Status still proper and desired15. ☐ Certified Copy of Priority Document(s)  
(if foreign priority is claimed)16. ☐ Other: \_\_\_\_\_

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:



Continuation



Divisional



Continuation-in-part (CIP)

of prior application No. \_\_\_\_\_

## 18. CORRESPONDENCE ADDRESS

☒

Customer Number or Bar Code Label

05514

(Insert Customer No. or Attach bar code label here)

or ☐

Correspondence address below

NAME

Address

City

Country

State

Telephone

Zip Code

Fax

| CLAIMS   | (1) FOR | (2) NUMBER FILED | (3) NUMBER EXTRA | (4) RATE                      | (5) CALCULATIONS |
|--|---------|------------------|------------------|-------------------------------|------------------|
| TOTAL CLAIMS<br>(37 CFR 1.16(e))   |         | 24-20 =          | 4                | X \$ 18.00 =                  | \$ 72.00         |
| INDEPENDENT CLAIMS<br>(37 CFR 1.16(b))                                     |         | 6-3 =            | 3                | X \$ 78.00 =                  | \$ 234.00        |
| MULTIPLE DEPENDENT CLAIMS (if applicable) (37 CFR 1.16(d))                 |         |                  |                  | \$280.00 =                    | \$ 00.00         |
|  |         |                  |                  | BASIC FEE<br>(37 CFR 1.16(e)) | \$ 760.00        |
| Total of above Calculations =  |         |                  |                  |                               | \$1066.00        |
| Reduction by 50% for filing by small entity (Note 37 CFR 1.9, 1.27, 1.28). |         |                  |                  |                               |                  |
| TOTAL =  |         |                  |                  |                               | \$1066.00        |

## 19. Small entity status

- a. ☐ A Small entity statement is enclosed
- b. ☐ A small entity statement was filed in the prior nonprovisional application and such status is still proper and desired.
- c. ☐ Is no longer claimed.

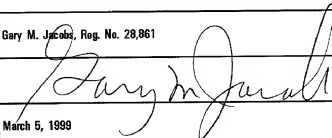
20. ☒ A check in the amount of \$ 1066.00 to cover the filing fee is enclosed.

21. ☐ A check in the amount of \$ \_\_\_\_\_ to cover the recordal fee is enclosed.

22. The Commissioner is hereby authorized to credit overpayments or charge the following fees to Deposit Account No. 06-1205:

- a. ☒ Fees required under 37 CFR 1.16.
- b. ☐ Fees required under 37 CFR 1.17.
- c. ☐ Fees required under 37 CFR 1.18.

## SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

|           |  |
|-----------|--|
| NAME      | Gary M. Jacobs, Reg. No. 28,861  |
| SIGNATURE |  |
| DATE      | March 5, 1999  |

TITLE OF THE INVENTION  
SPEECH SYNTHESIS APPARATUS, CONTROL METHOD THEREFOR,  
AND COMPUTER-READABLE MEMORY

5 BACKGROUND OF THE INVENTION

The present invention relates to a speech synthesis apparatus which has a database for managing phonemic piece data and performs speech synthesis by using the phonemic piece data managed by the database, a control method for the  
10 apparatus, and a computer-readable memory.

As a conventional speech synthesis method, a synthesis method based on a waveform concatenation scheme is available. In the waveform concatenation synthesis method, the prosody is changed by the pitch synchronous waveform overlap adding  
15 method of pasting waveform element pieces corresponding one to several pitches at desired pitch intervals. The waveform concatenation synthesis method can obtain more natural synthetic speech than a synthesis method based on a parametric scheme, but suffers the problem of a narrow  
20 allowable range with respect to changes in prosody.

Under the circumstances, attempts are made to improve the speech quality by preparing various speech data and properly selecting and using them. As a criterion for selection of speech data, information such as a phonemic  
25 context (a phoneme to be synthesized or a few phonemes on two sides of the target phoneme) or a fundamental frequency

F0 is used.

The following problems are, however, posed in the above conventional speech synthesis method.

If, for example, there is no data that satisfies a phonemic context as a synthesis target, a search for necessary speech data is made again by relaxing the condition associated with the phonemic context. The execution of this re-search in speech synthesis complicates the processing, resulting in an increase in processing time. In addition, when the fundamental frequency F0 is to be used as a criterion for selection of speech data, each speech data must be evaluated in association with the fundamental frequency F0 to obtain speech data that matches most with the fundamental frequency F0 of the speech data to be synthesized.

#### SUMMARY OF THE INVENTION

The present invention has been made in consideration of the above problems, and has as its object to provide a speech synthesis apparatus capable of performing speech synthesis with high precision at high speed, a control method therefor, and a computer-readable memory.

In order to achieve the above object, a speech synthesis apparatus according to the present invention has the following arrangement.

There is provided a speech synthesis apparatus having a database for managing phonemic piece data, comprising:

generating means for generating a second phoneme in consideration of a phonemic context for a first phoneme as a search target;

search means for searching the database for a phonemic  
5 piece data corresponding to the second phoneme;

re-search means for generating a third phoneme by changing the phonemic context on the basis of the search result obtained by the search means, and re-searching the database for phonemic piece data corresponding to the third  
10 phoneme; and

registration means for registering the search result obtained by the search means or the re-search means in a table in correspondence with the second or third phoneme.

In order to achieve the above object, a speech  
15 synthesis apparatus according the present invention has the following arrangement.

There is provided a speech synthesis apparatus for performing speech synthesis by using phonemic piece data managed by a database, comprising:

20 storage means for storing a table for managing position information indicating a position of phonemic piece data in the database in correspondence with a phoneme obtained in consideration of a phonemic context made to correspond to the phonemic piece data;

25 calculation means for acquiring each phonemic context information of a phoneme group as a synthesis target and

fundamental frequencies corresponding thereto and  
calculating an average of acquired fundamental frequencies;

search means for searching a phoneme group  
corresponding to the phonemic context information from the  
5 table;

acquisition means for acquiring, from the table,  
position information of phonemic piece data corresponding  
to a predetermined phoneme of the phoneme group searched out  
by the search means, on the basis of the average of fundamental  
10 frequencies calculated by the calculation means; and

changing means for acquiring phonemic piece data  
indicated by the position information acquired by the  
acquisition means from the database, and changing a prosody  
of the acquired phonemic piece data.

15 In order to achieve the above object, a control method  
for a speech synthesis apparatus according to the present  
invention has the following steps.

There is provided a control method for a speech  
synthesis apparatus having a database for managing phonemic  
20 piece data, comprising:

the generating step of generating a second phoneme in  
consideration of a phonemic context for a first phoneme as  
a search target;

the search step of searching the database for a  
25 phonemic piece data corresponding to the second phoneme;

the re-search step of generating a third phoneme by

changing the phonemic context on the basis of the search result obtained in the search step, and re-searching the database for phonemic piece data corresponding to the third phoneme; and

- 5           the registration step of registering the search result obtained in the search step or the re-search step in a table in correspondence with the second or third phoneme.

In order to achieve the above object, a control method for a speech synthesis apparatus according to the present  
10   invention has the following steps.

There is provided a control method for a speech synthesis apparatus for performing speech synthesis by using phonemic piece data managed by a database, comprising:

- 15           the storage step of storing a table for managing position information indicating a position of phonemic piece data in the database in correspondence with a phoneme obtained in consideration of a phonemic context made to correspond to the phonemic piece data;

- 20           the calculation step of acquiring each phonemic context information of a phoneme group as a synthesis target and fundamental frequencies corresponding thereto and calculating an average of acquired fundamental frequencies;

- 25           the search step of searching a phoneme group corresponding to the phonemic context information from the table;

          the acquisition step of acquiring, from the table,

position information of phonemic piece data corresponding to a predetermined phoneme of the phoneme group searched out in the search step, on the basis of the average of fundamental frequencies calculated in the calculation step; and

- 5       the changing step of acquiring phonemic piece data indicated by the position information acquired in the acquisition step from the database, and changing a prosody of the acquired phonemic piece data.

In order to achieve the above object, a  
10   computer-readable memory according to the present invention has the following program codes.

There is provided a computer-readable memory storing program codes for controlling a speech synthesis apparatus having a database for managing phonemic piece data,  
15   comprising:

      a program code for the generating step of generating a second phoneme in consideration of a phonemic context for a first phoneme as a search target;

      a program code for the search step of searching the  
20   database for a phonemic piece data corresponding to the second phoneme;

      a program code for the re-search step of generating a third phoneme by changing the phonemic context on the basis of the search result obtained in the search step, and  
25   re-searching the database for phonemic piece data corresponding to the third phoneme; and



a program code for the registration step of registering the search result obtained in the search step or the re-search step in a table in correspondence with the second or third phoneme.

- 5           In order to achieve the above object, a computer-readable memory according to the present invention has the following program codes.

There is provided a computer-readable memory storing program codes for controlling a speech synthesis apparatus  
10 for performing speech synthesis by using phonemic piece data managed by a database, comprising:

- a program code for the storage step of storing a table for managing position information indicating a position of phonemic piece data in the database in correspondence with  
15 a phoneme obtained in consideration of a phonemic context made to correspond to the phonemic piece data;

- a program code for the calculation step of acquiring each phonemic context information of a phoneme group as a synthesis target and fundamental frequencies corresponding  
20 thereto and calculating an average of acquired fundamental frequencies;

a program code for the search step of searching a phoneme group corresponding to the phonemic context information from the table;

- 25           a program code for the acquisition step of acquiring, from the table, position information of phonemic piece data

corresponding to a predetermined phoneme of the phoneme group searched out in the search step, on the basis of the average of fundamental frequencies calculated in the calculation step; and

- 5        a program code for the changing step of acquiring phonemic piece data indicated by the position information acquired in the acquisition step from the database, and changing a prosody of the acquired phonemic piece data.

According to the present invention described above,  
10       a speech synthesis apparatus capable of performing speech synthesis with high precision at high speed, a control method therefor, and a computer-readable memory can be provided.

Other features and advantages of the present invention will be apparent from the following description taken in  
15       conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

#### BRIEF DESCRIPTION OF THE DRAWINGS

20       Fig. 1 is a block diagram showing the arrangement of a speech synthesis apparatus according to the first embodiment of the present invention; ,

Fig. 2 is a flow chart showing search processing executed in the first embodiment of the present invention;

25       Fig. 3 is a view showing an index managed in the first embodiment of the present invention;

Fig. 4 is a flow chart showing speech synthesis processing executed in the first embodiment of the present invention;

Fig. 5 is a view showing a table obtained from the index managed in the first embodiment of the present invention;

Fig. 6 is a flow chart showing search processing executed in the second embodiment of the present invention;

Fig. 7 is a view showing an index managed in the second embodiment of the present invention:

Fig. 8 is a flow chart showing search processing executed in the third embodiment of the present invention; and

Fig. 9 is a view showing an index managed in the third embodiment of the present invention.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Fig. 1 is a block diagram showing the arrangement of a speech synthesis apparatus according to the first embodiment of the present invention.

Reference numeral 103 denotes a CPU for performing numerical operation/control, control on the respective components of the apparatus, and the like, which are executed in the present invention; 102, a RAM serving as a work area for processing executed in the present invention and a temporary saving area for various data; 101, a ROM storing various control programs such as programs executed in the

present invention, and having an area for storing a database 101a for managing phonemic piece data used for speech synthesis; 109, an external storage unit serving as an area for storing processed data; and 105, a D/A converter for  
5 converting the digital speech data synthesized by the speech synthesis apparatus into analog speech data and outputting it from a loudspeaker 110.

Reference numeral 106 denotes a display control unit for controlling a display 111 when the processing state and  
10 processing results of the speech synthesis apparatus, and a user interface are to be displayed; 107, an input control unit for recognizing key information input from a keyboard 112 and executing the designated processing; 108, a  
communication control unit for controlling  
15 transmission/reception of data through a communication network 113; and 104, a bus for connecting the respective components of the speech synthesis apparatus to each other.

Search processing of searching for a target phoneme, of the processing executed in the first embodiment, will be  
20 described next with reference to Fig. 2.

Fig. 2 is a flow chart showing search processing executed in the first embodiment of the present invention.

In the first embodiment, as phonemic contexts, two phonemes on both sides of each phoneme, i.e., phonemes as  
25 right and left phonemic contexts called a triphone, are used.

First of all, in step S1, a phoneme p as a search target

from the database 101a is initialized to a triphone ptr. In step S2, a search is made for the phoneme p from the database 101a. More specifically, a search is made for phonemic piece data having label p indicating the phoneme p. It is then

5 checked in step S4 whether there is the phoneme p in the database 101a. If it is determined that the phoneme p is not present (NO in step S4), the flow advances to step S3 to change the search target to a substitute phoneme having lower phonemic context dependency than the phoneme p. If the

10 phoneme p matching with the triphone ptr is not present in the database 101a, the phoneme p is changed to the right phonemic context dependent phoneme. If the right phonemic context dependent phoneme does not match with the triphone ptr, the phoneme p is changed to the left phonemic context

15 dependent phoneme. If the left phonemic context dependent phoneme does not match with the triphone ptr, the phoneme p is changed to another phoneme independently of a phonemic context. Alternatively, a high priority may be given to a left phonemic context phoneme for a vowel, and a high priority

20 may be given to a right phonemic context phoneme for a consonant. In addition, if there is no phoneme p that matches with the triphone ptr, one or both of left and right phonemic contexts may be replaced with similar phonemic contexts. For example, the "k" (consonant of the "ka" column in the Japanese

25 syllabary) may be used as a substitute when the right phonemic context is "p" (consonant for the "pa" column which is

modified "ha" column in the Japanese syllabary). Note, the Japanese syllabary is the Japanese basic phonetic character set. The character set can be arranged in a matrix where there are five (5) rows and ten (10) columns. The five rows are respectively the five vowels of the English language and the ten rows consist of 9 consonants and the column of the five vowels. A phonetic (sound) character is represented by the sound resulting from combining a column character and a row character, e.g. column "t" and row "e" is pronounced "te"; column "s" and row "o" is pronounced "so". After the phoneme p as the search condition is changed in this manner, the flow returns to step S2.

If it is determined that the phoneme p is present (YES in step S4), the flow advances to step S5 to calculate a mean F0 (the mean of the fundamental frequencies from the start of phonemic piece data to the end). Note that this calculation may be performed with respect to a logarithm F0 (function of time) or linear F0. Furthermore, the mean F0 of unvoiced speech may be set to 0 or estimated from the mean F0 of phonemic piece data of phonemes on both sides of the phoneme p by some method.

In step S6, the respective searched phonemic piece data are aligned (sorted) on the basis of the calculated mean F0. In step S7, the sorted phonemic piece data are registered in correspondence with the triphone ptr. As a result of registration, an index like the one shown in Fig. 3 is

obtained, which indicates the correspondence between generated phonemic piece data and triphones. As shown in Fig. 3, in the pointers managed in correspondence with the triphones, "phonemic piece position" indicating the location of each phonemic piece data in the database 101a and "mean F0" are managed in the form of a table.

Steps S1 to S7 are repeated for all conceivable triphones. It is then checked in step S8 whether the processing for all the triphones is complete. If it is determined that the processing is not complete (NO in step S8), the flow returns to step S1. If it is determined that the processing is complete (YES in step S8), the processing is terminated.

Speech synthesis processing of performing speech synthesis by searching for phonemic piece data of a phoneme as a synthesis target using the index generated by the processing described with reference to Fig. 2 will be described next with reference to Fig. 4.

Fig. 4 is a flow chart showing the speech synthesis processing executed in the first embodiment of the present invention.

When speech synthesis processing is to be performed, the triphone context ptr of the phoneme p as a synthesis target and F0 trajectory are given. Speech synthesis is then performed by searching phonemic piece data of phonemes on the basis of mean F0 and triphone context ptr and using the

waveform overlap adding method.

First of all, in step S9, mean  $F_0'$  which is mean of the given  $F_0$  trajectory of a synthesis target is calculated. In step S10, a table for managing the phonemic piece position of phonemic piece data corresponding to the triphone ptr of the phoneme p is searched out from the index shown in Fig. 3. If, for example, the triphone ptr is "a. A. b", the table shown in Fig. 5 is obtained from the index shown in Fig. 3. Since proper substitute phonemes have been obtained by the above search processing, the result of this step never becomes empty.

In step S11, the phonemic piece position of phonemic piece data having the mean  $F_0$  nearest to the mean  $F_0'$  is obtained on the basis of the table obtained in step S10. In this case, since the phonemic piece data have been sorted by the above search processing on the basis of mean  $F_0$ , a search can be made by using a binary search method or the like. In step S12, phonemic piece data is retrieved from the database 101a in accordance with the phonemic piece position obtained in step S11. In step S13, the prosody of the phonemic piece data obtained in step S12 is changed by using the waveform overlap adding method.

As described above, according to the first embodiment, when the absence of phonemic piece data is determined after the presence/absence of phonemic piece data is checked with respect to all the conceivable phonemic contexts, the



processing is simplified and the processing speed is increased by preparing substitute phonemes in advance. In addition, since information associated with the mean F0 of phonemic piece data present in each phonemic context is extracted in advance, and the phonemic piece data are managed on the basis of the extracted information. This can increase the processing speed of speech synthesis processing.

[Second Embodiment]

Quantization of the mean F0 of phonemic piece data may replace calculation of the mean F0 of continuous phonemic piece data in step S5 in Fig. 2 in the first embodiment. This processing will be described with reference to Fig. 6.

Fig. 6 is a flow chart showing search processing executed in the second embodiment of the present invention.

Note that the same step numbers in Fig. 6 denote the same processes as those in Fig. 2 in the first embodiment, and a detailed description thereof will be omitted.

In step S14, a mean F0 of the phonemic piece data of searched phonemes  $p$  is quantized to obtain the quantized mean F0 (obtained by quantizing the mean F0 as a continuous value at certain intervals). This calculation may be performed for the logarithm F0 or linear F0. In addition, the mean F0 of unvoiced speech may be set to 0, or unvoiced speech may be estimated from the mean F0 of phonemic piece data on both side of the unvoiced speech by some method.

In step S6a, the searched phonemic piece data are

aligned (sorted) on the basis of the quantized mean F0. In  
step S7a, the sorted phonemic piece data are registered in  
correspondence with triphones ptr. As a result of  
registration, an index indicating the correspondence between  
5 the generated phonemic piece data and the triphones is formed  
as shown in Fig. 7. In addition, as shown in Fig. 7, in the  
pointers managed in correspondence with the triphones,  
"phonemic piece position" indicating the location of each  
phonemic piece data in the database 101a and "mean F0" are  
10 managed in the form of a table.

Steps S1 to S7a are repeated for all possible triphones.  
It is then checked in step S8a whether the processing for  
all the triphones is complete. If it is determined that the  
processing is not complete (NO in step S8a), the flow returns  
15 to step S1. If it is determined that the processing is  
complete (YES in step S8a), the processing is terminated.

As described above, according to the second embodiment,  
in addition to the effects obtained in the first embodiment,  
the number of phonemic pieces and the calculation amount for  
20 search processing can be reduced by using the quantized mean  
F0 of phonemic piece data.

[Third Embodiment]

In the second embodiment, after the portions between  
the sorted phonemic piece data are interpolated, the  
25 respective phonemic piece data may be registered in  
correspondence with the triphones ptr. That is, an

arrangement may be made such that phonemic piece positions corresponding to the quantized means F0 of all the quantized phonemic piece data can be searched out in the tables in the index. This processing will be described with reference to

5 Fig. 8.

Fig. 8 is a flow chart showing search processing executed in the third embodiment of the present invention.

Note that the same step numbers in Fig. 8 denote the same processes as those in Fig. 6 in the second embodiment,  
10 and a detailed description thereof will be omitted.

In step S15, the portions between sorted phonemic piece data are interpolated. In step S7b, the interpolated phonemic piece data are registered in correspondence with triphones ptr. As a result of registration, an index  
15 indicating the correspondence between the generated phonemic piece data and the triphones is formed as shown in Fig. 9. In addition, as shown in Fig. 9, in the pointers managed in correspondence with the triphones, "phonemic piece position" indicating the location of each phonemic piece data in the  
20 database 101a and "mean F0" are managed in the form of a table.

Steps S1 to S7b are repeated for all possible triphones. It is then checked in step S8b whether the processing for all the triphones is complete. If it is determined that the processing is not complete (NO in step S8b), the flow returns  
25 to step S1. If it is determined that the processing is complete (YES in step S8b), the processing is terminated.

As described above, according to the third embodiment, in addition to the effects obtained in the second embodiment, since the phonemic piece positions of all phonemic piece data are managed, the processing in step S11 in Fig. 4 can be simply  
5 implemented as the step of referring to a table. This can further simplify the processing.

Note that the present invention may be applied to either a system constituted by a plurality of equipments (e.g., a host computer, an interface device, a reader, a  
10 printer, and the like), or an apparatus consisting of a single equipment (e.g., a copying machine, a facsimile apparatus, or the like).

The objects of the present invention are also achieved by supplying a storage medium, which records a program code  
15 of a software program that can realize the functions of the above-mentioned embodiments to the system or apparatus, and reading out and executing the program code stored in the storage medium by a computer (or a CPU or MPU) of the system or apparatus.

20 In this case, the program code itself read out from the storage medium realizes the functions of the above-mentioned embodiments, and the storage medium which stores the program code constitutes the present invention.

As the storage medium for supplying the program code,  
25 for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, CD-R, magnetic tape,

nonvolatile memory card, ROM, and the like may be used.

The functions of the above-mentioned embodiments may be realized not only by executing the readout program code by the computer but also by some or all of actual processing operations executed by an OS (operating system) running on the computer on the basis of an instruction of the program code.

Furthermore, the functions of the above-mentioned embodiments may be realized by some or all of actual processing operations executed by a CPU or the like arranged in a function extension board or a function extension unit, which is inserted in or connected to the computer, after the program code read out from the storage medium is written in a memory of the extension board or unit.

As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

WHAT IS CLAIMED IS

1. A speech synthesis apparatus having a database for managing phonemic piece data, comprising:

generating means for generating a second phoneme in  
5 consideration of a phonemic context for a first phoneme as a search target;

search means for searching said database for a phonemic piece data corresponding to the second phoneme;

re-search means for generating a third phoneme by  
10 changing the phonemic context on the basis of the search result obtained by said search means, and re-searching said database for phonemic piece data corresponding to the third phoneme; and

registration means for registering the search result  
15 obtained by said search means or said re-search means in a table in correspondence with the second or third phoneme.

2. The apparatus according to claim 1, wherein said registration means comprises

calculation means for calculating an average  
20 fundamental frequency of phonemic piece data searched out by said search means or said re-search means, and

sorting means for sorting the searched phonemic piece data group on the basis of the average fundamental frequency calculated by said calculation means, and

25 registers the phonemic piece data group and the second or third phoneme in correspondence with each other according

to an order in which the phonemic piece data group is sorted by said sorting means.

3. The apparatus according to claim 1, wherein the second phoneme is a triphone obtained in consideration of phonemic  
5 contexts of right and left phonemes of the first phoneme.

4. The apparatus according to claim 1, wherein the third phoneme is a phoneme obtained in consideration of at least one of phonemic contexts of right and left phonemes of the first phoneme.

10 5. The apparatus according to claim 1, wherein the third phoneme is a phoneme obtained in consideration of a left phonemic context of the first phoneme when the first phoneme is a vowel, and a right phonemic context of the first phoneme when the first phoneme is a consonant.

15 6. The apparatus according to claim 2, wherein said registration means further comprises quantization means for quantizing an average fundamental frequency of the searched phonemic piece data.

7. The apparatus according to claim 6, wherein said  
20 calculation means interpolates a frequency, of average fundamental frequencies of phonemic piece data groups quantized by said quantization means, for which no corresponding phonemic data is present by using an average fundamental frequency which is adjacent to the frequency and  
25 for which corresponding phonemic piece data is present.

8. A speech synthesis apparatus for performing speech

synthesis by using phonemic piece data managed by a database,  
comprising:

storage means for storing a table for managing position  
information indicating a position of phonemic piece data in  
5 the database in correspondence with a phoneme obtained in  
consideration of a phonemic context made to correspond to  
the phonemic piece data;

calculation means for acquiring phonemic context  
information of a phoneme as a synthesis target and  
10 fundamental frequencies corresponding thereto and  
calculating an average of acquired fundamental frequencies;

search means for searching a phoneme group  
corresponding to the phonemic context information from the  
table;

15 acquisition means for acquiring, from the table,  
position information of phonemic piece data corresponding  
to a predetermined phoneme of the phoneme group searched out  
by said search means, on the basis of the average of  
fundamental frequencies calculated by said calculation  
20 means; and

changing means for acquiring phonemic piece data  
indicated by the position information, acquired by said  
acquisition means from the database, and changing a prosody  
of the acquired phonemic piece data.

25 9. The apparatus according to claim 8, wherein said  
changing means changes the prosody by using a pitch



synchronous waveform overlap adding method.

10. The apparatus according to claim 8, wherein when a fundamental frequency of a phoneme obtained in consideration of the phonemic context is quantized, said storage means  
5 manages the quantized fundamental frequency in the table in correspondence with position information indicating a position in the database at which phonemic piece data corresponding to the phoneme is present.

11. The apparatus according to claim 8, wherein when a  
10 fundamental frequency of a phoneme obtained in consideration of the phonemic context is quantized, said calculation means acquires phonemic context information of a phoneme as a synthesis target, and calculates an average of quantized fundamental frequencies of the phoneme group.

15 12. A control method for a speech synthesis apparatus having a database for managing phonemic piece data, comprising:

a generating step of generating a second phoneme in consideration of a phonemic context for a first phoneme as  
20 a search target;

a search step of searching said database for a phonemic piece data corresponding to the second phoneme;

a re-search step of generating a third phoneme by changing the phonemic context on the basis of the search  
25 result obtained in said search step, and re-searching said database for phonemic piece data corresponding to the third

phoneme; and

a registration step of registering the search result obtained in said search step or said re-search step in a table in correspondence with the second or third phoneme.

- 5 13. The method according to claim 12, wherein said registration step comprises

a calculation step of calculating an average fundamental frequency of phonemic piece data searched out in said search step or said re-search step, and

- 10 a sorting step of sorting the searched phonemic piece data group on the basis of the average fundamental frequency calculated in said calculation step, and

- registering the phonemic piece data group and the second or third phoneme in correspondence with each other  
15 according to an order in which the phonemic piece data group is sorted in said sorting step.

14. The method according to claim 12, wherein the second phoneme is a triphone obtained in consideration of phonemic contexts of right and left phonemes of the first phoneme.

- 20 15. The method according to claim 12, wherein the third phoneme is a phoneme obtained in consideration of at least one of phonemic contexts of right and left phonemes of the first phoneme.

16. The method according to claim 12, wherein the third  
25 phoneme is a phoneme obtained in consideration of a left phonemic context of the first phoneme when the first phoneme

is a vowel, and a right phonemic context of the first phoneme when the first phoneme is a consonant.

17. The method according to claim 13, wherein said registration step further comprises a quantization step of  
5 quantizing an average fundamental frequency of the searched phonemic piece data.

18. The method according to claim 17, wherein said calculation step comprises interpolating a frequency, of  
average fundamental frequencies of phonemic piece data  
10 groups quantized in said quantization step, for which no corresponding phonemic data is present by using an average fundamental frequency which is adjacent to the frequency and for which corresponding phonemic piece data is present.

19. A control method for a speech synthesis apparatus for  
15 performing speech synthesis by using phonemic piece data managed by a database, comprising:

a storage step of storing a table for managing position information indicating a position of phonemic piece data in the database in correspondence with a phoneme obtained in  
20 consideration of a phonemic context made to correspond to the phonemic piece data;

a calculation step of acquiring phonemic context information of a phoneme as a synthesis target and fundamental frequencies corresponding thereto and  
25 calculating an average of acquired fundamental frequencies;

a search step of searching a phoneme group

corresponding to the phonemic context information from the table;

an acquisition step of acquiring, from the table, position information of phonemic piece data corresponding to a predetermined phoneme of the phoneme group searched out in the search step, on the basis of the average of fundamental frequencies calculated in said calculation step; and

a changing step of acquiring phonemic piece data indicated by the position information acquired in said acquisition step from the database, and changing a prosody of the acquired phonemic piece data.

20. The method according to claim 19, wherein said changing step comprises changing the prosody by using a pitch synchronous waveform overlap adding method.

21. The method according to claim 19, wherein when a fundamental frequency of a phoneme obtained in consideration of the phonemic context is quantized, said storage step comprises managing the quantized fundamental frequency in the table in correspondence with position information indicating a position in the database at which phonemic piece data corresponding to the phoneme is present.

22. The method according to claim 19, wherein when a fundamental frequency of a phoneme obtained in consideration of the phonemic context is quantized, said calculation step comprises acquiring phonemic context information of a phoneme as a synthesis target, and calculating an average

of quantized fundamental frequencies of the phoneme.

23. A computer-readable memory storing program codes for controlling a speech synthesis apparatus having a database for managing phonemic piece data, comprising:

5 a program code for the generating step of generating a second phoneme in consideration of a phonemic context for a first phoneme as a search target;

a program code for the search step of searching said database for a phonemic piece data corresponding to the  
10 second phoneme;

a program code for the re-search step of generating a third phoneme by changing the phonemic context on the basis of the search result obtained in the search step, and re-searching said database for phonemic piece data  
15 corresponding to the third phoneme; and

a program code for the registration step of registering the search result obtained in the search step or the re-search step in a table in correspondence with the second or third phoneme.

20 24. A computer-readable memory storing program codes for controlling a speech synthesis apparatus for performing speech synthesis by using phonemic piece data managed by a database, comprising:

a program code for the storage step of storing a table  
25 for managing position information indicating a position of phonemic piece data in the database in correspondence with

a phoneme obtained in consideration of a phonemic context made to correspond to the phonemic piece data;

5 a program code for the calculation step of acquiring phonemic context information of a phoneme as a synthesis target and fundamental frequencies corresponding thereto and calculating an average of acquired fundamental frequencies;

a program code for the search step of searching a phoneme group corresponding to the phonemic context information from the table;

10 a program code for the acquisition step of acquiring, from the table, position information of phonemic piece data corresponding to a predetermined phoneme of the phoneme group searched out in the search step, on the basis of the average of fundamental frequencies calculated in the calculation  
15 step; and

a program code for the changing step of acquiring phonemic piece data indicated by the position information acquired in the acquisition step from the database, and changing a prosody of the acquired phonemic piece data.

ABSTRACT OF THE DISCLOSURE

A second phoneme is generated in consideration of a phonemic context with respect to a first phoneme as a search target. Phonemic piece data corresponding to the second  
5 phoneme is searched out from a database. A third phoneme is generated by changing the phonemic context on the basis of the search result, and phonemic piece data corresponding to the third phoneme is re-searched out from the database. The search or re-search result is registered in a table in  
10 correspondence with the second or third phoneme.

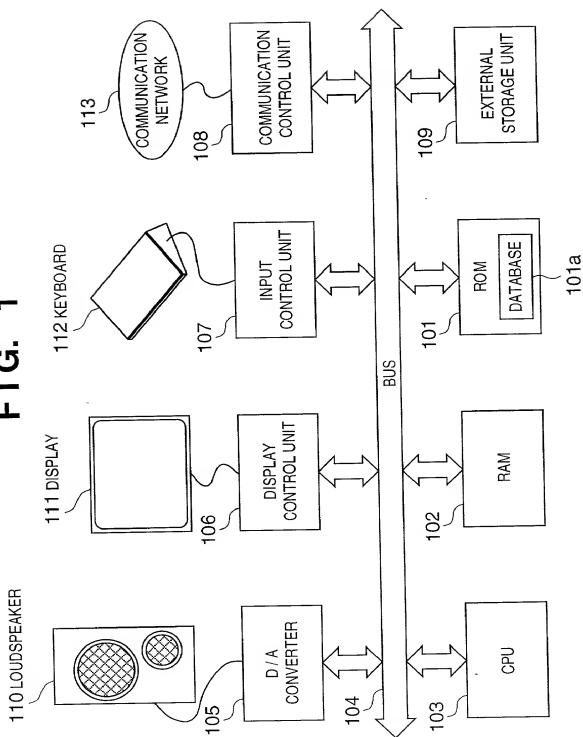
**FIG. 1**



FIG. 2

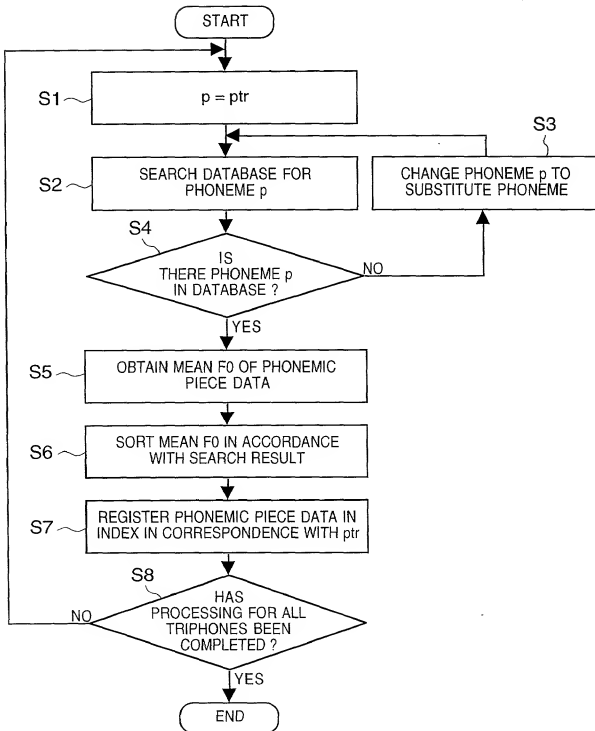




FIG. 4

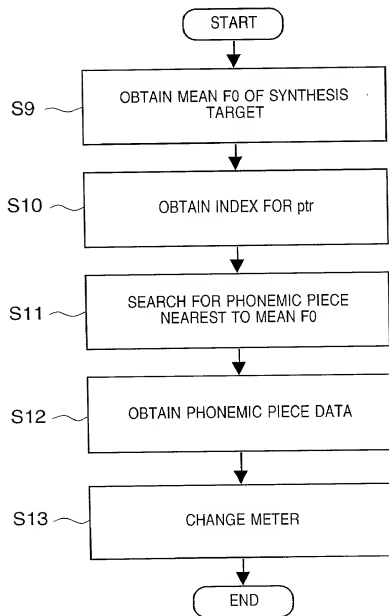


FIG. 5

|                            |       |       |       |     |       |       |
|----------------------------|-------|-------|-------|-----|-------|-------|
| MEAN F0                    | 98.3  | 130.7 | 240.4 | ... | 320.2 | 403.4 |
| PHONEMIC<br>PIECE POSITION | 14386 | 20056 | 23580 | ... | 30524 | 34216 |

FIG. 6

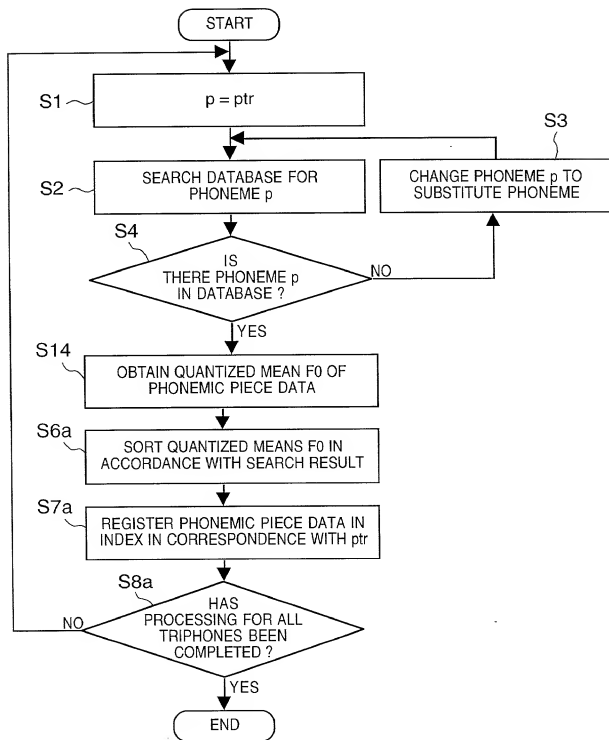


FIG. 7

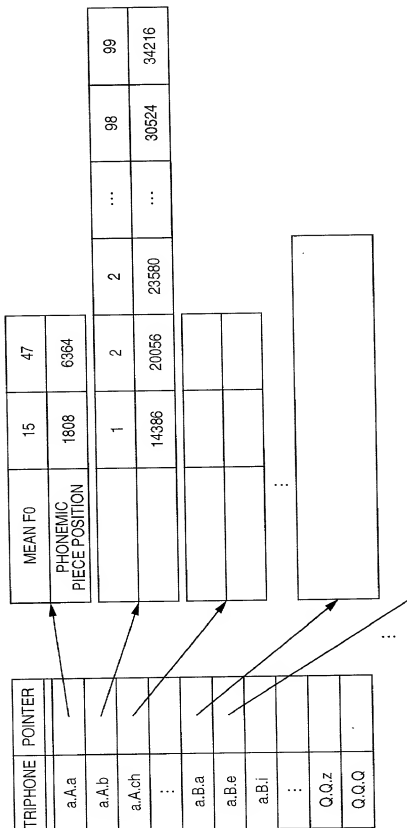


FIG. 8

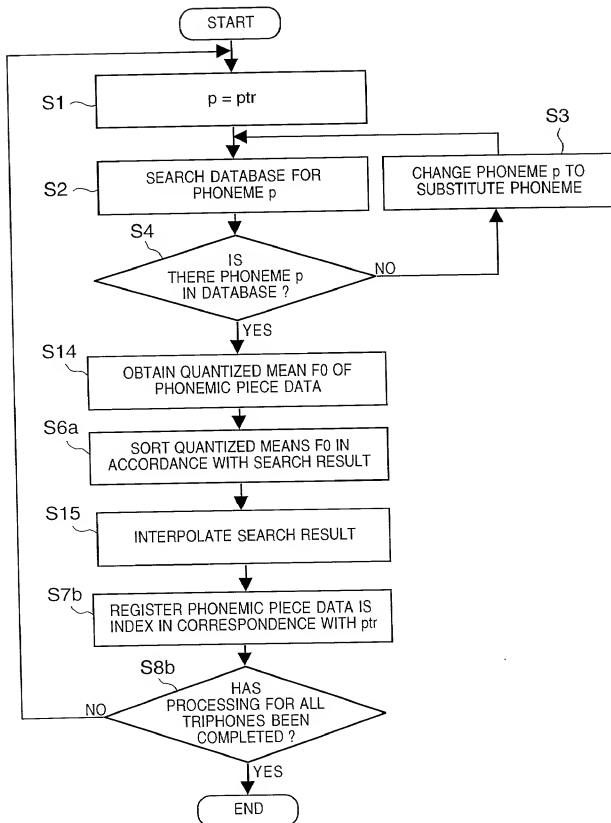


FIG. 9

| TRIPHONE | POINTER |
|----------|---------|
| a.A.a    |         |
| a.Ab     |         |
| a.A.ch   |         |
| :        |         |
| a.B.a    |         |
| a.B.e    |         |
| a.B.i    |         |
| :        |         |
| Q.Q.z    |         |
| Q.Q.Q    |         |

| MEAN F0                    | 1    | ... | 15   | ... | 47   | ... | 99   |
|----------------------------|------|-----|------|-----|------|-----|------|
| PHONEMIC<br>PIECE POSITION | 1808 | ... | 1808 | ... | 6364 | ... | 6364 |

|  |       |       |       |     |       |     |       |
|--|-------|-------|-------|-----|-------|-----|-------|
|  | 1     | 2     | 3     | ... | 98    | ... | 34216 |
|  | 14386 | 20056 | 23580 | ... | 30524 | ... |       |

|  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

|   |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|
| : |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|

|   |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|
| : |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|



**COMBINED DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION**  
(Page 1)

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name;

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

SPEECH SYNTHESIS APPARATUS, CONTROL METHOD THEREFOR, AND  
COMPUTER-READABLE MEMORY

the specification of which [ X ] is attached hereto. [ ] was filed on \_\_\_\_\_

as United States Application No. or PCT International Application No. \_\_\_\_\_  
and was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR §1.56.

I hereby claim foreign priority benefits under 35 U.S.C. §119(a)-(d) or §365(b), of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT international application which designates at least one country other than the United States, listed below and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed:

| <u>Country</u> | <u>Application No.</u> | <u>Filed (Day/Mo./Yr.)</u> | <u>(Yes/No)</u><br><u>Priority Claimed</u> |
|----------------|------------------------|----------------------------|--|
| JAPAN          | 10-057249              | 09/03/1998                 | Yes  |

I hereby appoint the practitioners associated with the firm and customer number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to the address associated with that Customer Number:

**FITZPATRICK, CELLA, HARPER & SCINTO**  
Customer Number: 05514

**COMBINED DECLARATION AND POWER OF ATTORNEY  
FOR PATENT APPLICATION**  
(Page 2)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Sole or First Inventor Masayuki YAMADA  
 Inventor's signature Masayuki Yamada  
 Date March 1, 1999 Citizen/Subject of Japan  
 Residence c/o Canon Miyazakidai-ryo, 1-3, Miyazaki 6-chome,  
Miyamae-ku, Kawasaki-shi, Kanagawa-ken, Japan  
 Post Office Address c/o CANON KABUSHIKI KAISHA,  
30-2, Shimomaruko 3-chome, Ohta-ku, Tokyo, Japan

Full Name of Second Joint Inventor, if any \_\_\_\_\_  
 Second Inventor's signature \_\_\_\_\_  
 Date \_\_\_\_\_ Citizen/Subject of \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Post Office Address \_\_\_\_\_

Full Name of Third Joint Inventor, if any \_\_\_\_\_  
 Third Inventor's signature : \_\_\_\_\_  
 Date \_\_\_\_\_ Citizen/Subject of \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Post Office Address: \_\_\_\_\_

Full Name of Fourth Joint Inventor, if any \_\_\_\_\_  
 Fourth Inventor's signature \_\_\_\_\_  
 Date \_\_\_\_\_ Citizen/Subject of \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Post Office Address \_\_\_\_\_

Full Name of Fifth Joint Inventor, if any \_\_\_\_\_  
 Fifth Inventor's signature \_\_\_\_\_  
 Date \_\_\_\_\_ Citizen/Subject of \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Post Office Address \_\_\_\_\_

Full Name of Sixth Joint Inventor, if any \_\_\_\_\_  
 Sixth Inventor's signature \_\_\_\_\_  
 Date \_\_\_\_\_ Citizen/Subject of \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Post Office Address \_\_\_\_\_